DOCUMENT RESUME

ED 255 575 TM 850 218

AUTHOR Frary, Robert B.; And Others
TITLE Partial Information and the "Correction" for Guessing.
PUB DATE Apr 85
NOTE 19p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, IL, April 1-3, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Guessing (Tests); Higher Education; *Multiple Choice Tests; Questionnaires; Scores; *Scoring Formulas; Testing Problems; Test Interpretation; Test Validity
IDENTIFIERS Instructions; *Test Score Variance

ABSTRACT

Students in an introductory college course (n=275) responded to equivalent 20-item halves of a test under number-right and formula-scoring instructions. Formula scores of those who omitted items averaged about one point lower than their comparable (formula adjusted) scores on the test half administered under number-right instructions. In contrast, average scores of students without omissions were nearly equal on the two halves. This interaction was statistically significant. This finding is contrasted with recently published results to the contrary, and reasons for this discrepancy are suggested. Support by measurement specialists for the use of formula scoring is analyzed in light of these findings and of misconceptions about formula scoring which are widespread among educators and the general public. It is concluded that the use of formula scoring to enhance the image of testing among this constituency is unethical. Appendices contain the written and oral instructions to the examinees and the questionnaire used to evaluate these instructions. (Author/DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED255575

Partial Information and the "Correction" for Guessing

Robert B. Frary, Lawrence H. Cross and Edward H. Sewell, Jr.

Virginia Polytechnic Institute and State University

Blacksburg, Virginia 24061

Abstract: Students in an introductory college course (N=275) responded to equivalent, 20-item halves of a test under number-right and formula-scoring instructions. Formula scores of those who omitted items averaged about one point lower than their comparable (formula adjusted) scores on the test half administered under number-right instructions. In contrast, average scores of students without omissions were nearly equal on the two halves. This interaction was statistically significant ($p < .02$). This finding is contrasted with recently published results to the contrary, and reasons for this discrepancy are suggested. Support by measurement specialists for use of formula scoring is analyzed in the light of these findings and misconceptions about formula scoring which are widespread among educators and the general public. It is concluded that use of formula scoring to enhance the image of testing among this constituancy is unethical.

Paper presented at the annual
meeting of the National Council
on Measurement in Education
Chicago, Illinois
April, 1985

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

✓ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. B. Frary

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 850 218

Partial Information and the "Correction" for Guessing

Robert B. Frary, Lawrence H. Cross and Edward H. Sewell, Jr.

Virginia Polytechnic Institute and State University

The arguments, pro and con, concerning the correction for guessing, or formula scoring, can be summed up briefly. In favor is its potential for reducing extraneous score variance due to guessing among all choices. Proponents point with alarm at the effect of this practice on the standard error of measurement of scores of low-ability examinees, who might guess among all choices on large proportions of items. Opponents claim that use of formula scoring introduces irrelevant sources of score variance due to personality factors and failure to understand the directions. In addition, opponents claim that formula scoring causes some examinees to suppress partial information (i.e., not to guess when it is to their advantage to do so) and is hence unfair.

This paper presents additional evidence against formula scoring, namely, that for a typical college examination it does suppress partial information and that misconceptions concerning it are widespread among examinees. The design of the study was adopted in answer to criticism of the design of earlier studies which showed suppression of partial knowledge under formula scoring. In these earlier studies (Slakter et al., 1976; Crehan et al., 1976; Cross & Frary, 1977; Poggio et al., 1978; and Bliss, 1980) examinees responded under formula scoring instructions and immediately afterward were asked to provide responses in place of their omissions. In all five studies, they answered correctly somewhat more often than expected under random guessing. However, reconsidering items after the fact is not the same as attempting them the first time, and, hence, these studies lack an element of realism necessary for greater generalizability of their results.

In contrast to these results, Angoff and Shrader (1984), using a design similar to that of the present study, concluded that supression of partial information under formula scoring was not a problem. In later discussion, we will present a possible reason for this divergence not only from recent outcomes but also from a long-term, consistent accumulation of evidence of undesirable aspects of formula scoring.[1] More important, however, is the fact that Angoff and Shrader also found no evidence to support use of formula scoring from the standpoint of its effect

---

[1]For example, see Rowley and Traub (1977), Slakter (1968a), Slakter (1968b), Mead and Smith (1957), Sherriffs and Boomer (1954), Votaw (1936), and Granich (1931). See also reviews by Abu-Sayf (1979) and Diamond and Evans (1973).

on reliability or validity. This outcome is consistent with research findings over several decades. Yet potential for improved reliability (and by extension, validity) is the backbone of the theoretical argument in favor of formula scoring.

One wonders then at the persistence of support for formula scoring. There is the issue of its possible beneficial effect on the psychometric characteristics of scores of low-ability examinees. If these were relatively few in number, the effect on overall reliability might not be discernable. However, the only evidence available about the responding of such examinees suggests that it is not generally random (see various reports on prediction from chance-level scores, e.g., Bolt, 1968).

The foregoing concerns notwithstanding, it hardly seems reasonable to penalize potentially large numbers of cautious examinees for the sake of the psychometric characteristics of the scores of a relative few at the bottom of the scale. This argument is made even stronger by the availability of alternative and more certain means for dealing with this problem (if it exists), for example, test lengthening by addition of easy items or provision of an alternative test for low scorers. This inconsistency led us to analyze reasons for support of formula scoring beyond fundamental psychometric considerations. We concluded that this support arises from misconceptions among educators and segments of the general public concerning partial information and its role in measurement. This analysis follows presentation of our empirical results.

Data Collection

The subjects were 275 students in an introductory course, Development of Communication, in the fall of 1984 at Virginia Tech, a comprehensive, moderately selective state university with 22,000 students. They were mostly freshmen and sophomores and represented a wide variety of majors. A 40-item test, one of two prior to the final examination, was split into halves (referred to as the X-half and Y-half) by the instructor on the basis of content equivalence. Four forms were then produced on which separate, written instructions preceded each half:

Form A - X-half with number-right instructions
Y-half with formula scoring instructions

Form B - Y-half with number-right instructions
X-half with formula scoring instructions

Form C - X-half with formula scoring instructions
Y-half with number-right instructions

Form D - Y-half with formula scoring instructions
X-half with number right instructions

Appendix A contains an explanation of the research, which was presented orally before the test, and the written instructions preceding each half. The four forms were passed out alternately prior to a 40-minute testing period, which was ample for all examinees. After this time, the test booklets were collected and a questionnaire (Appendix B) concerning issues underlying the research was passed out. Answers to the questionnaire were to be made on the same answer sheet used for the test. Participation at this point was optional, and 26 students (9%) declined.

Test Results

The first question to be answered was whether there were test-half or order-of-instructions effects. Accordingly, two 2x2 factorial analyses of variance (ANOVA) were performed, one on the number right scores and one on the formula scores. Independent variables were type of instructions and test half (two levels each). The largest F-ratio in these analyses had a probability of .22, and, indeed, corresponding cell means were nearly equal, as shown in Table 1. Therefore, all formula scores and number-right scores were combined for further analyses. Application of the Spearman-Brown formula to the values of coefficient alpha in Table 1 suggests total test reliability coefficients in the range of .58 to .75.

To permit comparisons between number-right and formula scores, each examinee's number-right score was corrected for guessing by application of the conventional formula. These scores will be referred to as corrected number-right (CNR) scores, not to be confused with the corrected-for-guessing (CFG) scores from the test halves administered under formula scoring directions.

The CFG and CNR scores were analyzed using a two-way, repeated-measures, factorial ANOVA. One factor was number of omissions on the formula-scored half of the test (four levels for 0, 1, 2, and 3 or more omissions). The second factor was response instruction, number-right or formula scoring; the repeated measures were under this factor, namely, subjects' CNR and CFG scores. Both factors and their interaction were significant ($p<.02$ for all three F-ratios). Table 2 shows the means for all cells of this design to illustrate the significant interaction. As shown there, higher numbers of omissions tend to be associated with higher CFG-CNR deficits. On the average, examinees with one or more omissions appear to have been "shortchanged" about one point on the 20-item test halves which were formula scored.

Questionnaire Results

Table 3 gives response percentages for the 18 questionnaire items. Items 1 through 4 were designed to evaluate the subjects' attention to the two sets of directions. Large proportions answered these questions correctly. However, 16% could not identify the correct formula scoring directions. In an earlier study using an item with the same wording, 27% answered incorrectly (Cross & Frary, 1977).

Items 5 through 7, 14, and 16 evaluate behavior under the two sets of instructions. Only 4% claimed more conservative behavior than the formula scoring instructions advised (No. 5). In contrast, 25% claimed more conservative behavior than recommended by number-right instructions (No. 6), and 19% claimed they would not mark unreached items scored number-right (No. 14). These tendencies did not come into play on the present test, however, as there were only four omissions among the potential 5500 number-right responses. While about half the sample claimed to act more conservatively under formula scoring (No. 7), those not claiming to do so included some of those reporting number-right strategies more conservative than recommended. About a third of examinees claimed generally to ignore directions (No. 16). This group included substantial proportions both of examinees who reported guessing contrary to formula scoring directions and of examinees who reported inappropriately conservative behavior under number-right directions.

Questions 8 through 13 and 15 elicited opinions and personal evaluations concerning test-taking. Substantial majorities considered themselves average or better in test-taking skills and luck on multiple-choice tests (Nos. 8 and 9). Another substantial majority associated the term "luck" with success in guessing as opposed to more indirect interpretations (No. 10). Not a majority, but 43%, believed that formula scoring penalizes examinees who guess at random (No. 11). Seventeen percent agreed or tended to agree that there was an element of dishonesty in guessing under any instructions (No. 12), and a substantial proportion of these subjects were among the 22% who considered formula scoring more fair than number-right (No. 13). Only 66% agreed or tended to agree that an examinee might ever find it expedient to guess among all choices hoping for short-term luck (No. 15). Those who disagreed or tended to disagree with this item included many of those who felt guessing was dishonest (No. 12), who claimed they would not answer unreached number-right items (No. 14), and who claimed behavior more conservative than recommended under number-right instructions (No. 6).

Except as noted above, interrelationships among responses to Nos. 1 through 16 were weak. Also, responses to Nos. 17 and 18, concerning gender and grade-point-average, were only minimally related to responses to any other questionnaire items.

The relationship between CFG-CNR differences and the questionnaire items was of considerable interest to us. Therefore, one-way ANOVA's were run on CFG-CNR using each of the questionnaire items as independent variables. Only two of these yielded F-ratios with probabilities of less than .05. These results suggest mean CFG-CNR deficits for those who claimed more conservative behavior under formula scoring (No. 7, p=.03) and, based on a post-hoc analysis of cell means, for those who disagreed that guessing among all choices might sometimes be advisable under formula scoring (No. 15, p=.04).

Discussion

The results just presented are consistent with most research concerning the use of formula scoring. They confirm suppression of partial information among those who omit and relate the resulting deficit to personal beliefs or attitudes. The size of the deficit is not large on the average. In fact, for some 40% of omitters, the CFG-CNR differences were positive. Probably some of the omitters in our study did not suppress partial information. Therefore, the effect of partial information suppression by the remainder is probably much more severe than the negative CFG-CNR mean over all omitters would indicate. We had hoped that some of the items on the questionnaire would help to isolate subgroups of omitters with more severe CFG-CNR deficits, but this outcome did not materialize. Nevertheless, to the extent that CFG-CNR was significantly related to questionnaire responses, it was observed that examinees with more conservative responses did tend to have larger deficits.

These outcomes should be compared to those of Angoff and Shrader (1984), which, as noted earlier, they interpreted to suggest that suppression of partial information does not occur to a meaningful extent when formula scoring is used. While the designs of these two studies were similar and, we believe, adequate, we would like to point out an unreconcilable problem with their data which may have masked the effect of formula scoring. This problem involves the presence of omissions under number-right scoring. Unlike our data, where these omissions were almost entirely absent, the Angoff and Shrader data include many omissions under number-right responding. Nor are all these omissions ones after the last responses (trailing as opposed to embedded omissions). For the various tests of their study, Angoff and Shrader report mean numbers of embedded omissions ranging from .44 to 4.74 under number-right responding. These are only means; the distributions are probably highly skewed, with many examinees having no embedded omissions. Thus for those who did make embedded omissions under number-right responding, the average number must have been substantial. Moreover, these are probably the same conservative or cautious individuals who suppress partial information under formula scoring. Had they been motivated to respond to the best of their ability, the mean CNR scores of the Angoff and Shrader study would probably have

been higher, quite possibly enough so to generate CFG-CNR differences similar to those we observed.[2]

The number-right omissions reported by Angoff and Shrader are quite in keeping with the behavior reported by subjects responding to our questionnaire. On one or more of Nos. 6, 12, 14, and 15, about 25% indicated reluctance to guess or disapproval of guessing even on tests scored number-right. The fact that they did not act on this belief on our test was probably due to the fact that they were familiar enough with the course material that they genuinely felt their guesses were justified. Nevertheless, when we imposed formula scoring, intimidation took over with the results reported above.

Analysis of Support for Formula Scoring

The foregoing theoretical and empirical considerations aside, measurement specialists who promote formula scoring are supported by some educators and segments of the general public who view score components from guessing (even guessing based on partial information) as ill-gotten gains. This constituency is not likely to disappear completely in the next generation if the questionnaire responses of our sample of students are any indication; 17% agreed or tended to agree that any guessing is dishonest (No. 12).

Some of these people simply fear the possibility that a lucky guesser may attain an inappropriately high score. However, they fail to understand that use of formula scoring prevents this outcome only to the extent that it suppresses guessing among all choices. The examinee who can eliminate at least one wrong choice on almost all items should guess to about the same extent under either formula or number-right scoring and will be rewarded relatively the same under either method.[3]

---

[2] If the omissions were replaced by random guesses, the CNR scores would average about the same as before. However, suppression of partial information even under number-right directions seems quite likely in view of the responses to questionnaire item No. 6 on which 25% reported more conservative behavior than the number-right instructions advised. Therefore, to the extent that partial information were applied in replacing the omissions, the resulting CNR average would be higher. In other words, the number-right averages would be higher in either case but more so if partial information were used in replacing omissions.

[3] This statement is based on the assumption that examinees guess whenever they can eliminate wrong choices. Then the expected formula scores are a linear function of the expected number-right scores, regardless of the extent to which the examinees guess among all choices (Frary, 1980). In other words, use of

In contrast, the examinee who knows little and desires or needs a high score has no recourse but to guess at every opportunity, even when formula scoring is used, hoping that luck will be favorable in the short run. To whatever extent luck is favorable, such an examinee will also be rewarded relatively the same under either method. The only examinees that formula scoring may prevent from attaining inappropriately high scores are those who are naturally compliant, those who fear the result of bad luck, and those who misunderstand the effect of formula scoring, believing it actually imposes a penalty for guessing.

Regardless of the preceding discussion, most people concerned about inappropriately high scores due to luck have little appreciation of the rarity of such occurrences. The probability of correctly guessing (at random) 21 or more out of 25 four-choice items is about one in a billion. For a more practical example, consider an examinee who knows five items and can eliminate all but two choices on each of the remaining 20. This examinee has an expected number right of 15, with chances of about one in 169 of answering 21 or more correctly and of about one in a million of making a perfect score.

Independent of concern about inappropriately high scores due to guessing is a desire on the part of some educators to have scores which reflect the number of items known without recourse to guessing. Usually this desire is associated with a (probably inappropriate) tendency to view the test in question as criterion-referenced and the scores as absolute measures. The formula scores are naively perceived as somehow purged of the components due to guessing (even guessing based on partial information). Davis (1967) actually investigated the extent to which formula scores approximate the number known without recourse to guessing and concluded that the adequacy of the approximation varied widely depending on the average number of wrong choices eliminated and the extent of misinformation (classification of the answer as a distractor).

In spite of the untenable nature of the beliefs just reviewed, some respected measurement specialists have actually supported formula scoring as a public relations measure (e.g., Stanley, 1954; Davis, 1964). Perhaps they believed the practice was, at worst, harmless. Whatever the motivation, use of formula scoring probably does generate a more favorable image of

---

formula scoring only rescales the score components attributable to assurred knowledge, partial information, random guessing, misinformation, etc., or, as expressed by Lord (1975), number-right and formula scores are both unbiased estimators of the same quantity. This fact should not be confused with that to the effect that observed formula scores are a linear function of observed number-right scores when there are no omissions.

multiple-choice testing among those who are misinformed or have unrealistic beliefs about guessing. Moreover, their lack of understanding would probably be difficult to overcome. Who then can blame measurement specialists who support formula scoring for basking in the approval of this misguided but possibly influential constituency?

We believe the time has come to answer this question publicly and unequivocally. We propose that it is simply immoral to continue to promote general use of formula scoring. There are special applications where it or some other approach to suppressing guessing may be appropriate, such as speeded tests and some criterion-referenced tests, especially for diagnostic purposes, where guessing might cause incorrect educational placement. However, general use of formula scoring on tests such as those of the Angoff and Shrader study cannot be justified and must be condemned on ethical grounds. Its use discriminates against a class of examinees that has been known to exist for almost as long as multiple-choice testing has been practiced, namely, the cautious, unsure individuals who are reticent to make choices which may be wrong. These individuals are not served completely fairly even by number-right testing, as evidenced by the number-right omissions reported by Angoff and Shrader.

Probably the only fair procedure is to inform examinees that choices will be made for them randomly in place of any omissions.[4] Then even the most cautious examinee might prefer to make a choice rather than leave the outcome completely to chance. This recommendation is almost the same as that by Cureton (1966) to the effect that formula scoring should be used whenever instructions fail to prevent omissions. However, this approach would not motivate those who would omit normally under number-right scoring and might actually encourage omissions.

In summary, we propose that:

1. Formula scoring offers no practical psychometric advantage.

2. Formula scoring tends to suppress partial information of cautious or compliant examinees to their disadvantage.

3. Given 1. and 2., the use of formula scoring for public relations purposes alone is unethical and should be neither recommended nor practiced.

---

[4]Alternatively, embedded omissions might be replaced with random responses, while trailing omissions would receive the score expected from random guessing.

REFERENCES

Abu-Sayf, F.K. (1979). The scoring of multiple-choice tests: A closer look. Educational Technology, 19, 5-15.

Angoff, W.H. and Schrader, W.B. (1984). A study of hypotheses basic to the use of rights and formula scores. Journal of Educational Measurement, 21, 1-17.

Bliss, L.B. (1980). A test of Lord's assumption regarding examinee guessing behavior on multiple-choice tests using elementary school students. Journal of Educational Measurement, 17 147-152.

Bolt, R.F. (1968). Study of linearily and homoscedasticity of test scores in the chance score range. Educational and Psychological Measurement, 28, 47-60.

Crehan, K.D., Candor, C.A., & Beckett, G.W. (1976). Utilization of partial knowledge on objective exams under formula and number-right scoring directions. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April.

Cross, L.H., and Frary, R.B. (1977). An empirical test of Lord's theoretical results regarding formula scoring of multiple-choice tests. Journal of Educational Measurement, 14, 313-321.

Cureton, E.E. (1966). The correction for guessing. Journal of Experimental Education, 4, 44-47.

Davis, F.B. (1964). Educational measurements and their interpretation. Belmont, California: Wadsworth.

Davis, F.B. (1967). A note on the correction for chance success. Journal of Experimental Education, 35, 42-47.

Diamond, J. and Evans, W. (1973). The correction for guessing. Review of Educational Research, 43, 181-191.

Granich, L.A. (1931). A technique for experimenting on guessing in objective tests. Journal of Educational Psychology, 22, 145-156.

Frary, R.B. (1980). The effect of misinformation, partial information and guessing on expected multiple-choice test item scores. Applied Psychological Measurement, 4, 79-90.

Lord, F.M. (1975). Formula scoring and number-right scoring. Journal of Educational Measurement, 12, 7-11.

Mead, A.R., & Smith, R.M. (1957). Does the true-false scoring formula work? Some data on an old subject. Journal of Educational Research, 51, 47-53.

Poggio, J.P., Asmus, E.P., and Levy, J. (1978). An investigation of responses to omitted items under formula scoring directions: Paper presented at the annual meeting of the American Educational Research Association, Toronto.

Rowley, G.L., & Traub, R.E. Formula scoring, number-right scoring, and test taking strategy. Journal of Educational Measurement, 14, 15-22.

Sherriffs, A.C., & Boomer, D.S. (1954). Who is penalized by the penalty for guessing? Journal of Educational Psychology, 45, 81-90.

Slakter, M.J. (1968a). The penalty for not guessing. Journal of Educational Measurement, 5, 141-144.

Slakter, M.J. (1968b). The effect of guessing strategy on objective test scores. Journal of Educational Measurement, 5, 217-221.

Slakter, M.J., Juliano, D.B., Treible, J., & Sarnacki, R.E. (1976). Is there a penalty for not guessing? Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April.

Stanley, J.C. (1954). Psychological correction for chance. Journal of Experimental Education, 22, 297-298.

Votaw, D.F. (1936). The effect of do-not-guess directions on the validity of true-false or multiple-choice tests. Journal of Educational Psychology, 27, 698-703.

Table 1 - Descriptive Test Score Statistics by Form

Form	Instruction Order	N	Mean	Std. Dev.	α
A	No. right	67	15.8	1.87	.46
	Formula		14.0	3.68	.60
B	No. right	71	16.0	2.40	.58
	Formula		14.5	2.65	.43
C	No. right	66	16.1	2.13	.52
	Formula		14.2	3.01	.48
D	No. right	71	15.7	2.06	.54
	Formula		14.6	3.11	.43

Table 2 - Means of CFG and CNR by Number of CFG Omissions

CFG Omissions	N	Mean CFG	Mean CNR	Mean CFG - CNR
0	177	15.09	14.92	.175
1	39	13.64	14.50	- .862
2	21	12.76	13.14	- .381
3 or more	19	11.53	13.00	-1.473

Table 3 - Response Percentages for Questionnaire Items
N = 249

Paraphrased Question	Responses	
1. Which test half was formula scored?	a. Right answer	94%
	b. Wrong answer	6%
2. Formula score is rights minus	a. number answered wrong	4%
	b. fraction of number wrong	90%
	c. number guessed wrong	1%
	d. fraction of number guessed wrong	4%
3. Formula directions were	a omit if uncertain	5%
	b. omit if only able to eliminate one or two choices	2%
	c. omit if guessing among all choices	84%
	d. not to omit any	8%
4. Number-right directions were	a. omit if uncertain	2%
	b. omit if only able to eliminate one or two choices	2%
	c. omit if guessing among all choices	6%
	d. not to omit any	90%
5. What did you do under formula scoring when uncertain?	a. Omitted	4%
	b. Guessed if able to eliminate or had a hunch	76%
	c. Guessed regardless	14%
	d. Was always certain	6%
6. What did you do under number right when uncertain?	a. Omitted	1%
	b. Guessed if able to eliminate or had a hunch	24%
	c. Guessed regardless	66%
	d. Was always certain	9%
7. Did you adopt a more conservative strategy under formula scoring?	a. Yes	49%
	b. No	51%
8. Rate you skill as a testaker.	a. Below average	12%
	b. Average	67%
	c. Above average	21%

9. How lucky are you on multiple-choice tests?	a. Rather unlucky	24%
	b. Neither lucky or unlucky	60%
	c. Rather lucky	16%
10. How did you interpret "lucky"?	a. Test questions matched what I studied.	19%
	b. Extent of correct guessing when unsure.	70%
	c. Background and talents had favorable influence.	10%
11. Correcting-for-guessing penalizes random guessers	a. True	43%
	b. False	57%
12. Element of dishonesty in guessing regardless of instructions.	a. Agree	5%
	b. Tend to agree	12%
	c. Tend to disagree	29%
	d. Disagree	55%
13. Is formula or number-right scoring more fair?	a. Corrected-for guessing	22%
	b. Number-right	30%
	c. Neither more fair	48%
14. If no time to finish, I would mark unreached questions	a. not at all	19%
	b. only if number-right used	61%
	c. even if correction used	20%
15. Sometimes desirable to guess at random under formula scoring hoping for short-term luck	a. Agree	22%
	b. Tend to agree	42%
	c. Tend to disagree	23%
	d. Disagree	12%
16. I generally ignore directions.	a. Agree	13%
	b. Tend to agree	21%
	c. Tend to disagree	25%
	d. Disagree	41%
17. Sex	a Male	38%
	b. Female	62%
18. QCA	a. less than 2.0	12%
	b. 2.0 - 2.49	26%
	c. 2.5 - 2.99	19%
	d. 3.0 - 3.49	8%
	e. 3.5 and up	8%
	f. beginning freshman	27%

APPENDIX A - Instructions to Examinees

Preliminary Oral Instructions

The test you are about to take was made up by Professor Sewell and consists of 40 multiple-choice questions. Four forms of the test have been prepared: Forms A, B, C, and D. The same 40 questions appear on each form. The only difference is the order of the questions.

Each form has two sets of directions, one for the first 20 questions and a different set for the last 20 questions. Not only are the directions different for the first and second halves, but each half will be scored differently. Therefore it is very important that you read both sets of directions carefully when you get your test questions.

Your score on this test will be the sum of your scores on the two halves of the test. Grades will be assigned to the scores as is the usual practice for this course.

You will have 40 minutes to complete both parts of the test. At three o'clock we will collect your test questions and distribute a brief questionnaire designed to assess your beliefs about multiple-choice tests. Please record your answers to these questions on the same machine-readable answer sheet you used for the test beginning with number 41.

Written Instructions Preceding Test Halves

Formula Scoring

Your score on this part of the test will be the number of questions answered correctly minus a fraction of the number you answer incorrectly:

SCORE = Number right - (1/3) Number wrong

It is to your advantage to answer questions even when you are not completely sure your answers are correct. This is especially true if you can eliminate one or more choices as incorrect or have a hunch or feeling about which choice is correct. However, it is better to omit a question than to guess wildly among all the choices given.

Number-Right Instructions

Your score on this part of the test will be the number of questions answered correctly. It is to your advantage, therefore, to answer every question in this part of the test, even if you should have to choose some answers at random.

APPENDIX B - Questionnaire Items

Directions: The questions which follow assess your understanding and reaction to the two sets of directions (number-right and formula scoring) provided in the test you just took. Please answer the questions to the best of your ability with the knowledge that your answers to these questions will not affect your grade in any way.

1. As you recall, formula scoring will be used to score which half of the test you took?

   a. First half
   b. Last half

2. According to the directions, your score on the formula scored part of the test will be computed as the number of questions answered correctly minus

   a. the number of questions answered incorrectly.
   b. a fraction of the questions answered incorrectly.
   c. the number of questions guessed incorrectly.
   d. a fraction of the questions guessed incorrectly.

3. The formula scoring directions advised you

   a. to omit answering a question if you were not certain of the answer.
   b. to omit answering a question if you could eliminate only one or two of the choices as incorrect.
   c. to omit answering a question if you had to guess among all choices.
   d. not to omit any questions under any condition.

4. The number-right scoring directions advised you

   a. to omit answering a question if you were not certain of the answer.
   b. to omit answering a question if you could eliminate only one or two of the choices as incorrect.
   c. to omit answering a question if you had to guess among all choices.
   d. not to omit any questions under any condition.

5. In the set of formula scored questions, what did you usually do when uncertain of a correct answer?

   a. Omitted the item.
   b. Guessed only if I had a hunch as to the correct answer or could eliminate at least one choice as incorrect.
   c. Guessed even if I had no idea which was the correct answer.
   d. Not applicable. I was certain of the correct answer to all questions scored this way.

6. In the set of questions scored number right, what did you usually do when uncertain of a correct answer?

   a. Omitted the item.
   b. Guessed only if I had a hunch as to the correct answer or could eliminate at least one choice as incorrect.
   c. Guessed even if I had no idea which was the correct answer.
   d. Not applicable. I was certain of the correct answer to all questions scored this way

7. For those questions to be formula scored, did you consciously adopt a more conservative guessing strategy than for those questions to be scored number-right?

   a. Yes
   b. No

8. In general, how would you rate your skill as a test taker?

   a. Below average
   b. About average
   c. Above average

9. In general, how lucky are you in taking multiple-choice tests?

   a. Rather unlucky
   b. Neither lucky nor unlucky
   c. Rather lucky

10. Thinking of your answer to the above question, which of the following best fits your interpretation of the word "lucky"?

    a. The extent to which the test questions happen to match what you studied.
    b. The extent to which you are successful in guessing the correct choice when unsure of an answer.
    c. Your background and talents provide a good basis for learning the tested materials.

11. Formula scoring is designed to penalize those who disregard the directions and make sheer guesses on multiple-choice tests.

    a. True
    b. False

12. There is an element of dishonesty in guessing when unsure of the answers to multiple-choice questions, regardless of how the test is scored.

    a. Agree
    b. Tend to agree
    c. Tend to disagree
    d. Disagree

13. Which method of scoring multiple-choice tests do you believe is more fair for determining student grades.

    a. Formula scoring
    b. Number-right scoring
    c. Neither is more fair

14. If I did not have time to read and consider some of the questions on a multiple-choice test, I would mark choices for these questions

    a. not at all.
    b. only if number-right scoring was to be used.
    c. even if formula scoring was to be used.

15. Even if formula scoring is to be used, circumstances might make it advisable for some examinees to guess among all choices in the hope that chance will be favorable in the short run.

    a. Agree
    b. Tend to agree
    c. Tend to disagree
    d. Disagree

16. When all is said and done, I tend to ignore the test directions and do what comes naturally when I take a multiple-choice test.

    a. Agree
    b. Tend to agree
    c. Tend to disagree
    d. Disagree

17. I am

    a. male.
    b. female.

18. My QCA is in the range

    a. less than 2.0
    b. 2.0 - 2.49
    c. 2.5 - 2.99
    d. 3.0 - 3.49
    e. 3.5 or higher
    f. Unknown (first quarter freshman)